

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/29024>

Please be advised that this information was generated on 2017-12-05 and may be subject to change.

Protein Sequences Indicate That Turtles Branched Off from the Amniote Tree After Mammals

Gert-Jan Caspers,¹ Geert-Jan Reinders,¹ Jack A.M. Leunissen,² Jan Wattel,³ Wilfried W. de Jong^{1,3}

¹Department of Biochemistry, University of Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

²CAOS/CAMM Center, University of Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

³Institute for Systematics and Population Biology, University of Amsterdam, P.O. Box 94766, 1090 GT Amsterdam, The Netherlands

Received: 6 October 1995 / Accepted: 24 January 1996

Abstract. The phylogenetic relationships among the major groups of amniote vertebrates remain a matter of controversy. Various alternatives for the position of the turtles have been proposed, branching off either before or after the mammals. To discover the phylogenetic position of turtles in relation to mammals and birds, we have determined cDNA sequences for the eye lens proteins α A- and α B-crystallin of the red-eared slider turtle (*Trachemys scripta elegans*). In addition, databases were searched for turtle protein sequences, for which mammalian, avian, and outgroup orthologs were available. All sequences were analyzed by three phylogenetic tree reconstruction methods (neighbor-joining, maximum parsimony, and maximum likelihood). Including the α -crystallins, 7 out of 12 proteins support a sister-group relation of turtles and birds with all 3 methods. For each of the other five proteins no topology was consistently preferred by the three approaches. Analyses of the combined amino acid data (1,695 aligned sites) also give extremely strong evidence that turtles are nearer to birds, indicating that mammals branched off before the divergence between turtles and birds occurred.

Key words: Testudines — *Trachemys scripta elegans* — Tetrapod phylogeny — Molecular evolution — α -crystallin

Introduction

The major groups of amniote vertebrates diverged from a common ancestor during a relatively short period at the end of the Paleozoic era, about 300–250 million years ago (Carroll 1987; Laurin and Reisz 1995). Until recently there was a broadly accepted view of the relationships between these amniote groups, based on paleontological and morphological evidence. This held, as already implied by Haeckel (1866), that mammals represent the sister group of all other extant amniotes. The turtles (Testudines) were generally considered to be the next group to have branched off, followed by Lepidosauria (tuatara, lizards, and snakes), with the final divergence occurring between crocodiles and birds (together the Archosauria) (Carroll 1987).

In the past decade this view has seriously been challenged with the advent of more sophisticated cladistic analyses. The hypothesis of a sister-group relation between mammals and birds (Gardiner 1982, 1993; Løvtrup 1985), reviving the clade Haematothermia (Owen 1866), has stirred much debate. Also, the position of the turtles is controversial. Some authors unite Testudines and Archosauria, to the exclusion of Lepidosauria (Løvtrup 1977; Hennig 1983; Ax 1984). Alternatively, the turtles have been proposed to be the first to branch off from the amniotes, followed by the mammals (Gaffney 1980). This latter view has indeed been adopted in some recent text books (e.g., Chaline 1990, p 78; Ridley 1993, p 465). Various authors have, however, provided renewed support for the classical branching pattern, based on morphological characters (Carroll 1987; Gauthier et

al. 1988; Benton 1990; Laurin and Reisz 1995). A total evidence approach, combining morphological and molecular data, also resulted in the traditional phylogeny (Eernisse and Kluge 1993). The bird–crocodilian sister-group relationship has further been confirmed by a recent exhaustive study of molecular sequence data (Hedges 1994). However, from a molecular point of view the position of the turtles relative to mammals and other amniotes has not been well studied. Only few turtle sequences have been used in phylogenetic analysis, and results from different molecules have been contradictory (Hedges et al. 1990; Marshall 1992; Eernisse and Kluge 1993; Van de Peer et al. 1993).

As a contribution to solving the phylogenetic position of the turtles, we determined cDNA sequences for the eye lens proteins α A- and α B-crystallin from a turtle and analyzed these sequences together with all informative turtle protein sequences retrieved from the databases. Because birds are the best-represented nonmammalian amniotes in the databases, we limited ourselves to resolving the four-taxon case ((turtles, mammals, birds) outgroup). This would allow us to decide whether turtles branched off from the amniote tree before or after the mammals. It would simultaneously contribute to further settling the Haemathermia controversy. While four-taxon approaches have been widely used in phylogenetic studies (e.g., Graur et al. 1991; Steel et al. 1993), one should realize that these can be misleading in certain cases, as pointed out by Philippe and Douzery (1994). Therefore, if available, sequences of more than one species were used for each of the four major clades. In addition to the α A- and α B-crystallin sequences, exhaustive database searches yielded sequences of ten proteins for which orthologs are available from turtles, mammals, birds, and outgroups. We applied the three major tree construction methods—neighbor-joining, maximum parsimony, and maximum likelihood—on all sequences, apart and combined.

Materials and Methods

Sequence Analysis of α -Crystallins. Total RNA was isolated from eyes of three juvenile red-eared slider turtles, *Trachemys scripta elegans*, by the lithium chloride/urea method (Auffray and Rougeon 1980). RNA was reverse transcribed using SuperScript reverse transcriptase (Gibco BRL/Life Technologies) and oligo(dT) primer. α -Crystallin sequences were amplified from the resulting single-stranded cDNA by the polymerase chain reaction (PCR) method, using *Taq* polymerase (Gibco BRL/Life Technologies). Degenerated oligonucleotide primers were designed to amplify cDNA sequences coding for amino acid positions 12–160 of α A-crystallin and positions 9–61 of α B-crystallin. For α A-crystallin, the primers were as described earlier (Caspers et al. 1994); for α B-crystallin we used 5'-ATACTGCAGGATATCACCATTCA-CACCC-3' and 5'-ATAAAGCTTACCTC[C,T]GAGAGTCCCG-T[A,C,G,T]TC-3'. Hybridization temperatures were 45°C for the α A-crystallin amplification and 55°C for the α B-crystallin amplification. Reamplification with the same primers was necessary to obtain enough amplification product. Amplification products were made blunt-ended and 5'-phosphorylated with Klenow DNA polymerase and T4 poly-

nucleotide kinase (Gibco BRL/Life Technologies), respectively, and ligated into pGEM-3Zf(+) phagemid vector (Promega), cut with *Sma*I. Constructs from several separate amplifications were sequenced in both directions using the Sequenase 2.0 DNA sequencing kit (United States Biochemical). The sequences have been deposited in the GenBank data base (accession nos. U31938 and U31939).

Data Base Searches. The SwissProt (version 31.0), PIR (version 44.0), EMBL (version 42.0), and GenBank (version 83.0) databases were searched for amino acid sequences and protein-coding nucleic acid sequences from turtle species. Only sequences for which at least a mammalian, avian, and outgroup homolog were available were used. Sequences suspected not to obey the orthology criterion (i.e., resulting from a gene duplication rather than a speciation event) were excluded (e.g., gastrin/cholecystokinin, parvalbumin, and ribonuclease). However, as long as we found no evidence against the hypothesis that homologous sequences diverged from one another close in time to the ancestral divergence of their species lineages, they were treated as orthologs (Goodman et al. 1987); in cases where two different protein sequences of one species were available (β -hemoglobin, insulin), both were included in the analyses. Where several turtle sequences were known, species from all available families were included. If available, the mammalian sequences used were human, bovine, mouse, and a marsupial. For birds, chicken, and, if available, another species, preferably a paleognath, were used. When possible, multiple outgroups were used. Species names and database accession numbers are given in the legends of Fig. 2 and Table 1.

Phylogenetic Analyses. Because of the divergence times dealt with in this study, we used amino acid rather than nucleotide sequences for phylogenetic analyses. Sequences were aligned with PILEUP from the GCG package (Devereux et al. 1984). Phylogenetic analyses were performed with maximum parsimony (PROTPARS), neighbor-joining (Saitou and Nei 1987) (PROTDIST and NEIGHBOR, or TREECON) using Kimura distances (Kimura 1983), and maximum-likelihood methods (Felsenstein 1981) (PROTML) based on the JTT model (Jones et al. 1992). The programs PROTPARS, PROTDIST, and NEIGHBOR are from the PHYLIP package (Felsenstein 1993), TREECON has been written by Van de Peer and De Wachter (1994) and PROTML by Adachi and Hasegawa (1992). Confidence in the maximum-parsimony and neighbor-joining methods was assessed by bootstrapping (Felsenstein 1985) (SEQBOOT and CONSENSE (Felsenstein 1993) for the programs from the PHYLIP package). Confidence in the maximum-likelihood method was assessed by the differences in log-likelihood from the highest-likelihood tree (Bishop and Friday 1985).

Results

α -Crystallin Sequences

α -Crystallins belong to the small heat-shock protein family (Caspers et al. 1995). They occur in the vertebrate eye lens as multimeric complexes composed of two types of homologous subunits, α A- and α B-crystallin (Groenen et al. 1994). Both subunits are encoded by single-copy genes (King and Piatigorsky 1983; Quax-Jeuken et al. 1985), which avoids the problem of paralogy in comparative studies. α A-Crystallin protein sequences have already contributed to resolving the relationships between mammals, lizards, crocodiles, and birds (Stapel et

B

Fig. 1. Comparison of variable sites in turtle and other vertebrate α A-crystallins (**A**) and α B-crystallins (**B**). Deduced amino acid sequences of turtle *Trachemys scripta elegans* α A- and α B-crystallin cDNAs (positions 12–160 and 9–61, respectively) were aligned with corresponding sequences of α -crystallins from other species, selected to cover the variety within extant vertebrates. Only those positions are

shown at which different residues occur in these data sets. Amino acid position numbers are those of the complete α -crystallin chains. *Dots* indicate where residues are identical to the turtle sequences; *dashes* denote deletions. Species names and database accession numbers are given in the legends of Fig. 2 and Table 1.

Sequences from Databases

To combine the phylogenetic information from these newly determined α -crystallin sequences with that from other proteins available in the databases, exhaustive searches were performed for turtle sequences that might enable the resolution of the turtle–mammal–bird relationship. Amino acid sequences from ten additional sets of orthologous genes were found to be suitable for this purpose (Table 1). Selection criteria are described in Materials and Methods.

Each data set was subjected to maximum parsimony, neighbor-joining, and maximum-likelihood analyses. The support for the three alternative branching orders of mammals (M), turtles (T), and birds (B) with respect to outgroup sequences is summarized in Table 1. This table overwhelmingly demonstrates the evidence for a (M(T,B)) topology. Seven out of 12 protein sequences support a sister-group relation of turtle and birds in all analyses. Prolactin and nicotinamide adenine dinucleotide dehydrogenase subunit 2 (ND2) support a (M(T,B)) topology with two of the methods, while cytochrome *b*, myoglobin, and somatotropin support a (T(B,M)) topology with maximum parsimony and maximum likelihood. In the case of prolactin the likelihood support is equal for all three topologies.

Combining all proteins in a single analysis provides the strongest measure for the resolution of the four-taxon case under investigation. To that end a composite out-group sequence was constructed from the phylogenetically nearest outgroups (see footnote f of Table 1). For the ingroups, a turtle, chicken, human, and mouse were the only taxa available for all proteins. The combined analyses of all amino acid sequences (1,695 aligned sites) supports the sister-group relationship of turtle and birds almost at the highest possible levels (Table 1, Fig. 2B).

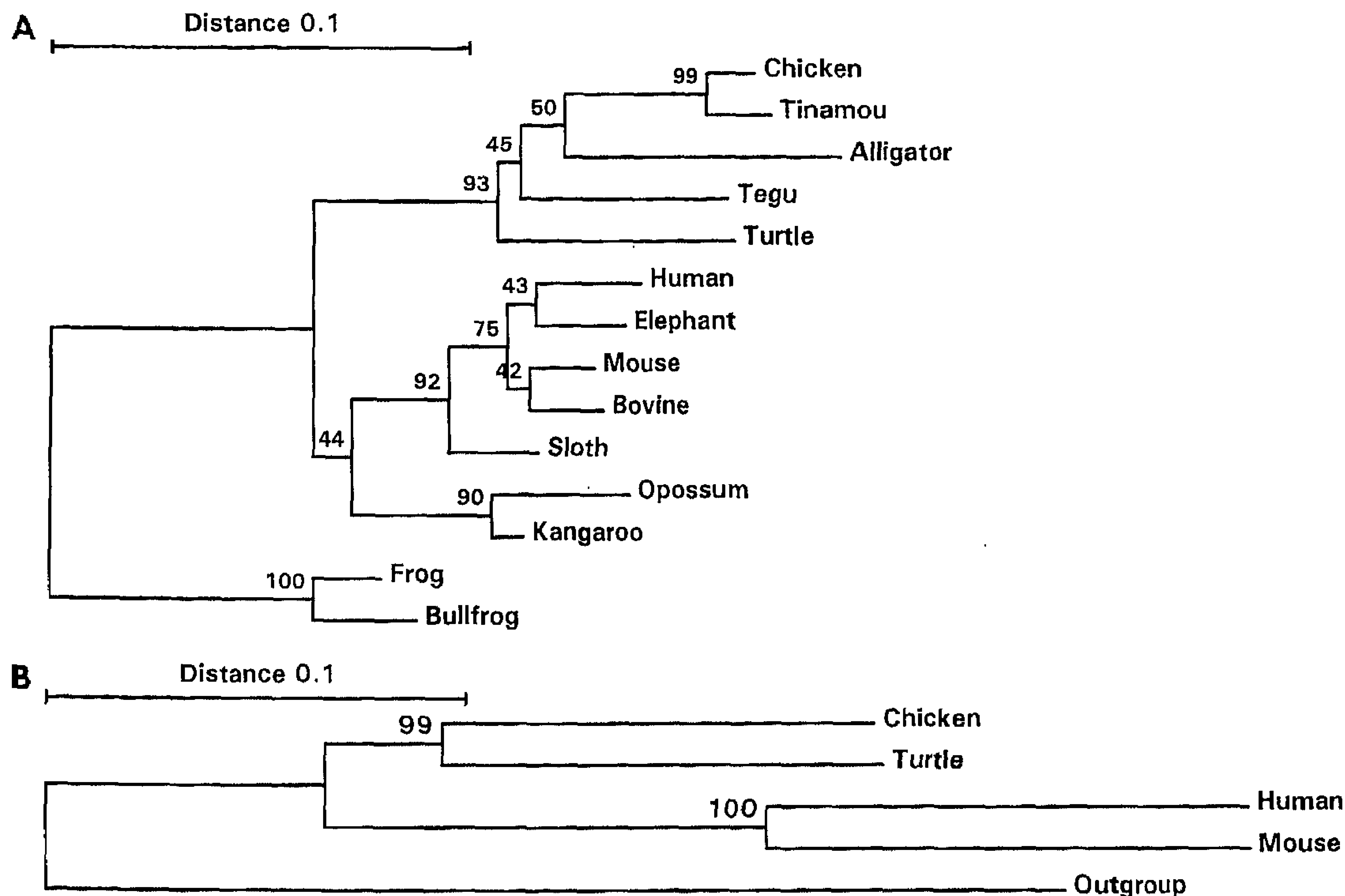


Fig. 2. Amniote relationships inferred from (A) α A-crystallin amino acid sequences (see Fig. 1A), and (B) the 12 combined protein sequences from Table 1. Neighbor-joining trees are shown, constructed with TREECON (Van de Peer and De Wachter 1994) using Kimura distances, with bootstrap values from 1,000 replications. Distance is proportional in the minimum number of mutations per residue. With the combined protein sequences, the fact that branch lengths in the reptile–bird lineage are considerably shorter than the branches leading to the mammalian species, also observed in the maximum-likelihood tree (not shown), is not due to a general acceleration in the evolutionary rate in the mammalian lineage but rather to particular more divergent sequences (mouse insulin and human cytochrome *c* and somatotropin).

Species names and database accession numbers for α A-crystallins are: chicken *Gallus gallus* (P02504), tinamou *Eudromia elegans* (L25850), alligator *Alligator mississippiensis* (P06904), tegu lizard *Tupinambis teguixin* (P02506), human *Homo sapiens* (P02489 with minor correction according to L25781), elephant *Loxodonta africana* (P02498), mouse *Mus musculus* (P02490), bovine *Bos taurus* (P02470), sloth *Choloepus hoffmanni* (P02486), opossum *Didelphis virginiana* (P02503), kangaroo *Macropus rufus* (P02502), frog *Rana esculenta* (up to amino acid position 70)/*R. temporaria* (from position 71 onward) (P02507 and P02508), and bullfrog *R. catesbeiana* (X85205). For species names and database accession numbers used in the combined protein analysis, see legend of Table 1.

Discussion

Previous studies of protein and nucleic acid sequences failed to give conclusive evidence about the position of turtles among the amniotes. 18S rRNA placed turtles outside a bird–mammal clade (Hedges et al. 1990; Eernisse and Kluge 1993), although weighting the nucleotide positions generated the classical amniote tree (Van de Peer et al. 1993), or nearly so (Marshall 1992). 28S rRNA sequences, which had earlier been inconclusive (Hedges et al. 1990), did group turtles within a bird–reptile clade, but as sister group to crocodiles, while salamanders were also included in the bird–reptile clade in this analysis (Eernisse and Kluge 1993). Studies of turtle insulin (Cascone et al. 1991), prolactin (Yasuda et al. 1990), somatotropin (Yasuda et al. 1989), and tyrosinase (Yamamoto et al. 1992; Morrison et al. 1994) noted that these proteins were closer to avian than to mammalian orthologs. α - and β -hemoglobin sequences, and a combined analysis of α - and β -hemoglobin, myoglobin, and cytochrome *c* placed turtles in a clade with croco-

diles and birds, excluding a mammal–lepidosaur clade (Eernisse and Kluge 1993). In earlier studies, β -hemoglobin and myoglobin grouped birds with mammals, both in maximum-parsimony (Goodman et al. 1987; Hedges et al. 1990) and in maximum-likelihood analyses (Bishop and Friday 1988). However, different taxonomic sampling may have played a role in these deviating results. Finally, mitochondrial tRNA sequences contained little phylogenetic information for inferring the position of turtles among the amniotes, while flanking protein-coding sequences (ND2) supported a placement of turtles as a sister group to a bird–crocodile clade (Seutin et al. 1994).

It might be expected that, like in the case of the avian sister-group controversy (Hedges 1994), an extended set of various proteins will enable a more convincing resolution of the branching order of turtles, mammals, and birds within the amniotes. We therefore determined sequences of two additional turtle proteins, α A- and α B-crystallin. Database searches yielded a further ten proteins suitable for our purpose. We analyzed these se-

Table 1. Protein sequence evidence for the branching order of mammals (M), turtles (T), and birds (B), with respect to outgroup sequences, as inferred by maximum-parsimony, neighbor-joining, and maximum-likelihood analyses^a

Protein ^c	Number of sites			Bootstrap support for alternative sister-group relationships in maximum parsimony analyses ^b		
	Alignable	Variable	Informative	(M(T,B))	(T(B,M))	(B(T,M))
αA-Crystallin*	149	63	7	92 (89)	3 (+3)	5 (+3)
αB-Crystallin*	53	33	4	94 (61)	<1 (+7)	6 (+5)
Cytochrome <i>b</i>	167	64	6	29 (+0)	41 (155)	16 (+1)
Cytochrome <i>c</i> *	104	22	2	69 (41)	26 (+1)	3 (+2)
α-Hemoglobin*	142	99	10	72 (321)	16 (+0)	2 (+2)
β-Hemoglobin*	146	117	19	67 (468)	15 (+1)	12 (+4)
Insulin*	51	14	3	91 (25)	4 (+3)	3 (+3)
Myoglobin	153	123	17	14 (+4)	85 (403)	1 (+6)
ND2	60	53	2	45 (255)	25 (+0)	7 (+2)
Prolactin	203	170	3	41 (525)	7 (+0)	37 (−2)
Somatotropin	193	119	9	30 (+0)	42 (226)	28 (+1)
Tyrosinase*	277	125	9	58 (249)	16 (+3)	25 (+2)
Combined proteins ^f	1695	850	64	98 (1,728)	1 (+25)	<1 (+29)

Protein ^c	Bootstrap support for alternative sister-group relationships in neighbor-joining analyses ^c			Differences in log-likelihood from best trees for alternative sister-group relationships in maximum likelihood analyses ^d		
	(M(T,B))	(T(B,M))	(B(T,M))	(M(T,B))	(T(B,M))	(B(T,M))
αA-Crystallin*	97	3	<1	ML	−1.9 ± 2.2	−1.9 ± 2.2
αB-Crystallin*	99	<1	<1	ML	−11.6 ± 7.0	−11.6 ± 7.0
Cytochrome <i>b</i>	24	29	36	−1.8 ± 3.8	ML	−2.7 ± 3.1
Cytochrome <i>c</i> *	44	42	7	ML	−1.3 ± 3.6	−2.4 ± 2.7
α-Hemoglobin*	94	4	<1	ML	−5.4 ± 4.0	−5.4 ± 4.0
β-Hemoglobin*	87	9	<1	ML	−0.3 ± 8.6	−7.8 ± 5.2
Insulin*	100	<1	<1	ML	−32.5 ± 1.6	−32.5 ± 1.6
Myoglobin	50	50	<1	−2.1 ± 2.6	ML	−1.6 ± 3.0
ND2	20	42	23	ML	−2.0 ± 2.2	−0.8 ± 3.4
Prolactin	85	3	12	ML	ML	ML
Somatotropin	71	19	10	−0.1 ± 4.2	ML	−2.0 ± 2.4
Tyrosinase*	50	22	29	ML	−1.4 ± 2.2	−1.5 ± 2.1
Combined proteins ^f	99	1	<1	ML	−38.7 ± 16.5	−51.4 ± 14.6

^a Values in boldface type emphasize the typology supported by the three different methods of tree construction. Asterisks (*) indicate the proteins that support a sister-group relationship of turtle and bird in all analyses

^b Bootstrap values in % (1,000 replications) that support the respective branching orders are given, followed (in parentheses) by the number of steps in the majority-rule consensus parsimony tree or the number of extra steps required for the alternative sister-group relationships. These are not necessarily the numbers of steps in the actual most parsimonious trees. The numbers of informative sites refer to the branching orders of mammals, turtles, and birds, not to branching orders within any of these clades

^c Bootstrap values in % (1,000 replications) that support the respective branching orders are given

^d The highest likelihood trees are indicated by ML. The differences of log-likelihoods, with their SEs, are given for the alternative topologies

^e Species names and database accession numbers for the proteins used in the analyses are: for αA-crystallin as given in the legend of Fig. 2, excluding alligator, tegu, elephant, and sloth, in order to obtain a similar data set as for the other proteins; for αB-crystallin, human (P02511), bovine (P02510), mouse (P23927), chicken (Q05713), duck *Anas platyrhynchos* (Q05557), and bullfrog (X87114); cytochrome *b*, green turtle *Chelonia mydas* (L12716), leatherback turtle *Dermochelys coriacea* (L12712), snapping turtle *Chelydra serpentina* (L12713), human (P00156), bovine (P00157), mouse (P00158), opossum *Monodelphis domestica* (Q04911), chicken (P18946), titmouse *Parus inornatus* (P29638), and clawed toad *Xenopus laevis* (P00160); cytochrome *c*, snapping turtle *C. serpentina* (P00022), human (P00001), bovine (P00006), mouse (P00009), kangaroo *M. giganteus* (P00014), chicken (P00016), ostrich *Struthio camelus* (P0001g), and bullfrog (P00024); α-hemoglobin, Western painted turtle *Chrysemys picta bellii* (P13273), human P01922), bovine (P01966), mouse (P01942), opossum *D. virginiana* (P01976), chicken (P01944), ostrich (P01981), clawed toad *X. tropicalis* (P07428), and axolotl *Ambystoma mexicanum* (P02015); β-hemoglobin, Western painted turtle (P13274), human (P02023), bovine (P02070), mouse (P02088 and P02089), opossum *D. virginiana* (P02109), chicken (P02112), ostrich (P02123), clawed toad *X. laevis* (P02132 and P02133), and newt *Triturus*

cristatus (P10785 and P10786); insulin, slider turtles *Trachemys scripta* and *Chrysemys dorsigni* mouse (P01325 and P01326), opossum *D. virginiana* (P18109), chicken and ostrich (P01332), and clawed toad *X. laevis* (P12706 and P12707); myoglobin, green turtle (P02202), map turtle *Graptemys geographica* (P02201), human (P02144), bovine (P02192), mouse (P04247), opossum *D. virginiana* (P02193), chicken (P02197), penguin *Aptenodytes forsteri* (P02199), carp *Cyprinus carpio* (P02204), tuna *Thunnus albacares* (P02205), and shark *Galeorhinus australis* (P14397); nicotinamide adenine dinucleotide dehydrogenase subunit 2 (ND2; including 5–6 residues of cytochrome *c* oxydase subunit I), loggerhead turtle *Caretta caretta* (Seutin et al. 1994), terrapin turtle *Maldemys terrapin* (Seutin et al. 1994), human (P03891/P00395), bovine (P03892/P00396), mouse (P03893/P00397), opossum *D. virginiana* (P41305/P41310), chicken (P18937/P18943), quail *Coturnix japonica* (P24971/P24984), clawed toad *X. laevis* (P03894/P00398), and carp (P24972/P24985); prolactin, green turtle (P33090), human (P01236), bovine (P01239), mouse (P06879), lungfish *Protopterus aethiopicus* (P33091), carp (P09585), and eel *Anguilla anguilla* (P33096); somatotropin, green turtle (P34005), human (P01241), bovine (P01246), mouse (P06880), chicken (P08998), duck (P11228), and bullfrog (P10813); tyrosinase, snapping turtle *Trionyx sinensis* (S56789), human (P14679), mouse (P11344), chicken (S54189), quail (S56788), and frog *R. nigromaculata* (Q04604). Due to computational limitations, in the maximum-likelihood analyses one of the mouse (P02089), toad (P02133), and newt (P10786) sequences from the β-hemoglobin data set and tuna from the myoglobin data set had to be removed

^f For the combined protein analysis, the turtle (green turtle for cytochrome *b* and myoglobin, loggerhead for ND2), chicken, human, and mouse (P02088 and P01325, respectively, in the cases of the recent duplications of the hemoglobin β and insulin genes) sequences were used. A composite outgroup sequence was constructed from the nearest available outgroups: frog αA-crystallin, bullfrog αB-crystallin, clawed toad cytochrome *b*, frog cytochrome *c*, toad α-hemoglobin, toad β-hemoglobin (P02132), toad insulin (P12706), carp myoglobin, toad ND2, lungfish prolactin, frog somatotropin, and frog tyrosinase

quences with three conceptually different methods of phylogeny reconstruction. Applying all three methods—neighbor-joining, maximum parsimony, and maximum likelihood—avoids discussions about their respective merits (Swofford and Olsen 1990; Hasegawa and Fujiwara 1993; Hillis et al. 1993, 1994; Tatenko et al. 1994; DeBry and Abele 1995; Huelsenbeck 1995). Observed congruence among branching patterns derived from independent data sets and by different methods provides the strongest obtainable evidence for phylogenetic robustness.

It is clear from Table 1 that a sister-group relation of turtles and birds, to the exclusion of mammals, is extremely well supported, especially when the sequences are combined. Unfortunately, the presently available protein data sets are not yet sufficient to completely resolve the branching order of the amniotes. The position of the lepidosaur sequence in the α A-crystallin trees is very weakly supported. Other proteins grouped lepidosaurs with mammals, which does not correspond to any morphology-based phylogeny (Eernisse and Kluge 1993). However, molecular analyses, combining amino acid and nucleic acid data, have already firmly established the sister-group relationship of crocodiles and birds (Hedges 1994). The present analysis further refutes the Haematothermia hypothesis and restores the position of the turtles as having branched off after mammals. It is quite likely that additional molecular data will further confirm the classical view of amniote phylogeny as based on morphological and paleontological evidence.

Acknowledgments. This work was carried out under the auspices of the Netherlands Foundation for Life Sciences (SLW) with financial aid from the Netherlands Organization for Scientific Research (NWO).

References

- Adachi J, Hasegawa M (1992) Computer science monographs, no. 27: MOLPHY: programs for molecular phylogenetics. I. — PROTML: maximum likelihood inference of protein phylogeny. Institute of Statistical Mathematics, Tokyo
- Auffray C, Rougeon F (1980) Purification of mouse immunoglobulin heavy chain mRNAs from total myeloma tumor RNA. *Eur J Biochem* 107:303–314
- Ax P (1984) Das phylogenetische System. G Fischer Verlag, Stuttgart, pp 101–114
- Benton MJ (1990) Phylogeny of the major tetrapod groups: morphological data and divergence dates. *J Mol Evol* 30:409–424
- Bishop MJ, Friday AE (1985) Evolutionary trees from nucleic acid and protein sequences. *Proc R Soc B* 226:271–302
- Bishop MJ, Friday AE (1988) Estimating the interrelationships of tetrapod groups on the basis of molecular sequence data. In: Benton MJ (ed) *The phylogeny and classification of the tetrapods. Vol 1: amphibians, reptiles, birds*. Clarendon Press, Oxford, pp 33–58
- Carroll RL (1987) *Vertebrate paleontology and evolution*. WH Freeman, New York, pp 1–15, 192–216
- Cascone O, Turyn D, Dellacha JM, Machado VLA, Marques M, Vita N, Cassan C, Ferrara P, Guillemot JC (1991) Isolation, purification, and primary structure of insulin from the turtle *Chrysemys dorsibigni*. *Gen Comp Endocrinol* 84:355–359
- Caspers G-J, Leunissen JAM, de Jong WW (1995) The expanding small heat-shock protein family, and structure predictions of the conserved “ α -crystallin domain.” *J Mol Evol* 40:238–248
- Caspers G-J, Wattel J, de Jong WW (1994) α A-Crystallin sequences group tinamou with ratites. *Mol Biol Evol* 11:711–713
- Chaline J (1990) *Paleontology of vertebrates*. Springer Verlag, Berlin
- DeBry RW, Abele LG (1995) The relationship between parsimony and maximum-likelihood analyses: tree scores and confidence estimates for three real data sets. *Mol Biol Evol* 12:291–297
- de Jong WW, Zweers A, Versteeg M, Dessauer HC, Goodman M (1995) α -Crystallin A sequences of *Alligator mississippiensis* and the lizard *Tupinambis teguixin*: molecular evolution and reptilian phylogeny. *Mol Biol Evol* 2:484–493
- Devereux J, Haeberli P, Smithies O (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12: 387–395
- Eernisse DJ, Kluge AG (1993) Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Mol Biol Evol* 10:1170–1195
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368–376
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
- Felsenstein J (1993) PHYLIP: phylogeny inference package, version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle, WA
- Gaffney ES (1980) Phylogenetic relationship of the major groups of amniotes. In: Panchen AL (ed) *The terrestrial environment and the origin of the land vertebrates*. Academic Press, London, pp 593–610
- Gardiner BG (1982) Tetrapod classification. *Zool J Linn Soc* 74:207–232
- Gardiner BG (1993) Haematothermia: warm-blooded amniotes. *Cladistics* 9:369–395
- Gauthier J, Kluge AG, Rowe T (1988) Amniote phylogeny and the importance of fossils. *Cladistics* 4:105–209
- Goodman M, Miyamoto MM, Czelusniak J (1987) Pattern and process in vertebrate phylogeny revealed by coevolution of molecules and morphologies. In: Patterson C (ed) *Molecules and morphology in evolution: conflict or compromise?* Cambridge University Press, Cambridge, pp 141–176
- Graur D, Hide WA, Li W-H (1991) Is the guinea-pig a rodent? *Nature* 351:649–652
- Groenen PJTA, Merck KB, de Jong WW, Bloemendal H (1994) Structure and modifications of the junior chaperone α -crystallin: from lens transparency to molecular pathology. *Eur J Biochem* 225:1–19
- Haeckel E (1866) *Generelle Morphologie der Organismen*. G Reimer, Berlin, Tafel XII
- Hasegawa M, Fujiwara M (1993) Relative efficiencies of the maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. *Mol Phylog Evol* 2:1–5
- Hedges SB (1994) Molecular evidence for the origin of birds. *Proc Natl Acad Sci USA* 91:2621–2624
- Hedges SB, Moberg KD, Maxson LR (1990) Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Mol Biol Evol* 7:607–633
- Hennig W (1983) Testudines. In: Hennig W (ed) *Stammesgeschichte der Chordaten*. P Parey, Hamburg, pp 132–139
- Hillis DM, Allard MW, Miyamoto MM (1993) Analysis of DNA sequence data: phylogenetic inference. *Methods Enzymol* 224:456–487
- Hillis DM, Huelsenbeck JP, Cunningham CW (1994) Application and accuracy of molecular phylogenies. *Science* 264:671–677
- Huelsenbeck JP (1995) The robustness of two phylogenetic methods: four-taxon simulations reveal a slight superiority of maximum likelihood over neighbor joining. *Mol Biol Evol* 12:843–849
- Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of

- mutation data matrices from protein sequences. *Comput Appl Biosci* 8:275–282
- Kimura M (1983) *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge, p 75
- King CR, Piatigorsky J (1983) Alternative RNA splicing of the murine α A crystallin gene: protein-coding information within an intron. *Cell* 32:707–712
- Laurin M, Reisz RR (1995) A reevaluation of early amniote phylogeny. *Zool J Linn Soc* 113:165–223
- Lee DC, Kim RY, Wistow GJ (1993) An avian α B-crystallin: non-lens expression and sequence similarities with both small (HSP27) and large (HSP70) heat shock proteins. *J Mol Biol* 232:1221–1226
- Løvtrup S (1977) *The phylogeny of the Vertebrata*. J Wiley & Sons, London, pp 180–184
- Løvtrup S (1985) On the classification of the taxon Tetrapoda. *Syst Zool* 34:463–470
- Lu S-F, Pan F-M, Chiou S-H (1995) Sequence analysis of frog α B-crystallin cDNA: sequence homology and evolutionary comparison of α A, α B and heat shock proteins. *Biochem Biophys Res Commun* 216:881–891
- Marshall CR (1992) Substitution bias, weighted parsimony, and amniote phylogeny as inferred from 18S rRNA sequences. *Mol Biol Evol* 9:370–373
- Morrison R, Mason K, Frost-Mason S (1994) A cladistic analysis of the evolutionary relationships of the members of the tyrosinase gene family using sequence data. *Pigment Cell Res* 7:388–393
- Owen R (1866) *On the anatomy of vertebrates*, vol I. Longmans, Green, London, p 7
- Philippe H, Douzery E (1994) The pitfalls of molecular phylogeny based on four species, as illustrated by the Cetacea/Artiodactyla relationships. *J Mamm Evol* 2:133–152
- Quax-Jeuken Y, Quax W, van Rens G, Meera Khan P, Bloemendal H (1985) Complete structure of the α B-crystallin gene: conservation of the exon-intron distribution in the two nonlinked α -crystallin genes. *Proc Natl Acad Sci USA* 82:5819–5823
- Ridley M (1993) *Evolution*. Blackwell Scientific Publications, Boston
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Seutin G, Lang BF, Mindell DP, Morais R (1994) Evolution of the WANCY region in amniote mitochondrial DNA. *Mol Biol Evol* 11:329–340
- Stapel SO, Leunissen JAM, Versteeg M, Wattel J, de Jong WW (1984) Ratites as oldest offshoot of avian stem—evidence from α -crystallin A sequences. *Nature* 311:257–259
- Steel MA, Lockhart PJ, Penny D (1993) Confidence in evolutionary trees from biological sequence data. *Nature* 364:440–442
- Swofford DL, Olsen GJ (1990) Phylogeny reconstruction. In: Hillis DM, Moritz C (eds) *Molecular systematics*. Sinauer, Sunderland, MA, pp 411–501
- Tateno Y, Takezaki N, Nei M (1994) Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. *Mol Biol Evol* 11:261–277
- Van de Peer Y, Neefs J-M, De Rijk P, De Wachter R (1993) Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. *J Mol Evol* 37:221–232
- Van de Peer Y, De Wachter R (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Biosci* 10:569–570
- Yamamoto H, Kudo T, Masuko N, Miura H, Sato S, Tanaka M, Tanaka S, Takeuchi S, Shibahara S, Takeuchi T (1992) Phylogeny of regulatory regions of vertebrate tyrosinase genes. *Pigment Cell Res* 5:284–294
- Yasuda A, Kawauchi H, Papkoff H (1990) The complete amino acid sequence of prolactin from the sea turtle (*Chelonia mydas*). *Gen Comp Endocrinol* 80:363–371
- Yasuda A, Yamaguchi K, Papkoff H, Yokoo Y, Kawauchi H (1989) The complete amino acid sequence of growth hormone from the sea turtle (*Chelonia mydas*). *Gen Comp Endocrinol* 73:242–251